


Rapport d'évaluation du mémoire de thèse / Evaluation report of the PhD thesis

Doctorant PhD student	Nom prénom / Full name Ecole Doctorale / Doctoral School Titre thèse / PhD Title	Javier Olivares Romero Physics Bayesian Hierarchical Modelling of Young Stellar Clusters
Rapporteur Reviewer	Nom prénom / Full name Etablissement / Institution Statut, fonction / Status, position	Timothy Naylor University of Exeter Norman Lockyer Professor of Astrophysics

Qualité du mémoire, rédaction & illustrations / Thesis quality, style & illustrations


 Satisfaisant / Satisfactory [☒] Bon / Good [☐] Très bon / Very good [☐] Exceptionnel [☐]

Commentaires/comments :

The majority of the thesis is written in a way which is clear, though there are numerous spelling and grammatical errors. Only occasionally do these obscure meaning, and I have attempted to find all these in the report that follows. The layout is logical, though I do feel a better overview of the whole process could have been provided. The figures are, on the whole good with informative captions, though a few (Figures 4.3 and the left panel of Figure 4.9) are too crowded to allow the reader to extract all the information. The candidate is to be congratulated on the layout of the mathematics (amongst the best I have seen), though a table of symbols would have helped the reader, and avoided unnecessary re-use of symbols.


Contexte, état de l'art, collaborations / Background, state of the art, collaborations :

Commentaires/comments :

The application of Bayesian analysis to problems in stellar physics is still in its infancy (in contrast to the position in cosmology), and hence any study of the efficacy of these methods is timely. This analysis builds on that of Sarro et al (2014, A&A 563 A45) which is applied to slightly earlier version of the dataset used here by Bouy et al (2015, A&A 577 A148).

This thesis is part of a wider collaboration and benefits enormously from the data collected by the DANCe collaboration.

Qualité scientifique, méthodologie, expérimentations, validation Scientific quality, methodology, experiments, validation


 Satisfaisant / Satisfactory [☐] Bon / Good [☒] Très bon / Very good [☐] Exceptionnel [☐]

Commentaires/comments :

In discussing the scientific quality of a thesis of this nature, I think one has to make a clear distinction between a thesis which presents new data and a thesis such as this where the real intellectual effort is developing new analytical methods. In the former case, one can reasonably ask for an uncontroversial data analysis which makes the conclusions from the data clear. In this case, there is bound to be controversy about the techniques, and the scientific quality lies making good (though not necessarily unassailable) choices in the analysis methods, and also in being able to make clear the advantages and disadvantages of the techniques developed. Finally, one has to be able to validate the outcomes by comparison with previous measurements and at least understand the differences.

Against those criteria the scientific quality is good. I cannot find any examples of algorithms which have been developed which are demonstrably wrong, and there are validations against previous measurements.

My criticism (see later in the report) lies around quantifying some of the suggested reasons for the discrepancies with previous methods, and sometimes there could be a better discussion of alternative methods which might have been employed.

Apports personnels, originalité, valorisation, perspectives
Personal contributions, originality, valorization, prospects

Commentaires/comments :

The analysis presented here is a significant and original improvement on previous efforts, in particular the development of the luminosity function estimator. This, and many of the other developments should be applicable to the Gaia data that will become available in DR2 and later. Thus, it is well-placed to make an impact on the field.

Conclusions du rapporteur / Reviewer's conclusions

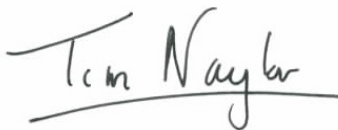
Commentaires/comments :

Although I present a long list of corrections in the full report below, none of them are take away from what I see as the main achievement of the thesis, which is the creation and exploration of a Bayesian hierarchical analysis which leads to the mass function. Given that, if the corrections can be made it would seem reasonable to proceed to the defence.

Avis du rapporteur / Reviewer's opinion :

Défavorable à la soutenance / Unfavorable to the defence ☐

Favorable ☒



Date 15 September 2017

Signature

Visa du directeur de l'école doctorale :

Rapport détaillé, commentaires libres, questionnements, correction demandées
Detailed report, free comments, questions, requested corrections.

Overview.

I have greatly enjoyed reading this thesis, and look forward to discussing it with the candidate. Overall the thesis is at the leading edge of a growing trend in modern stellar astronomy, where creating and understanding sophisticated statistical analysis is crucial.

The thesis begins in Chapter 1 by describing the stellar initial mass function, and some of its importance in astrophysics, and then moves on to discuss classifiers and Bayesian hierarchical models, and some of their applications in modern astrophysics. This literature review has an appropriate balance between statistical material and review of the science appropriate to the goals. Although the material reviewed could have been expanded, the size and depth is good enough given that the thesis has to discuss both science background and statistical methods,

Chapter 2 discusses previous work on the Pleiades, which is necessary background for the results which follow. As I state below I feel that in places a more critical analysis is required, but the important background is covered.

Chapter 3 is the main part of the thesis and describes how a Bayesian hierarchical model can be built for the data to hand. As I stated above, this is an extension of work within the group he is collaborating with, but also represents a significant advance in the area. Chapter 4 then presents the results of the analysis, placing it within the context of previous work, and Chapter 5 discusses what could be done in the future.

There are (inevitably) many things which I would like to discuss, and things which I think need changing. None of them, however, strike at the heart of what has been achieved, which I believe is the exploration of a new analysis pipeline.

Philosophical Issues.

In the introduction to probability theory (Page 45) the thesis states that it “is commonly agreed that the uncertainty of a measurement can be expressed in a probabilistic basis (JCGM 2008). It means that whenever we measure a quantity, a for example, then the distribution of the repeated measurements of a , follows a probability distribution function, $p(a)$.” This is a very surprising definition in a Bayesian-based thesis. In Bayesian statistics $p(a)$ represents your “reasonable degree of belief” (or perhaps understanding) of the possible values of a as a result of your experiment (see, for example, Section 1.2 of Gregory, 2005, “Bayesian Logical Data Analysis for the Physical Sciences”). If you have done your analysis correctly then I agree that repeating the experiment many times should give a distribution of results which corresponds to $p(a)$. This needs re-writing to reflect the fact that in Bayesian analysis no probability is simply a frequency (to paraphrase Jeffreys in Section VIII of “Theory of Probability”).

At the end of Section 1.4 (page 15) the candidate lays out what he intends to achieve. One of these is “Uncertainty of membership probabilities”. I do not understand what is meant here. A membership probability should represent the level of belief the author has that a star is a member of the cluster. As it is a binary choice, I do not believe it should have uncertainties. This is in contrast to estimating a parameter, where there is a most likely value, but the level of belief in alternative values must also be specified. Again, something different should be written here, and where this is referred to later in the thesis.

I think the final paragraph of Page 53 is wrong (though am quite happy for the candidate to argue with me). I agree that in some sense as the number of parameters increases the probability density must get lower per dimension. However, in assessing the evidence for a model we integrate over the space, so the density per dimension is not relevant.

Literature review (Chapters 1 and 2).

There are a few places where I disagree on points of fact, but my main requests for changes here center around enhancing the critical assessment of the material presented, sometimes with a greater emphasis on being quantitative.

Critical assessment of the literature.

The first of these examples is on Page 27, where there is a remark that Loktin (2006) failed to find any mass segregation, which is apparently in direct contradiction with the results of Pinfield (1998) discussed earlier, who does find it. A second example is the discussion of the luminosity function (Section 2.5), where the thesis presents a series of such distributions from historical sources, but no attempt is made to say which are right, and which later proved to be wrong. A final example is where the method of Limber (1960) for deriving an age is described, but there is no discussion as to whether it is technically correct, or why it gives the wrong age. All these three examples need a discussion which states what is right and what wrong, along with reasons.

There are two examples of where the discussion could be more quantitative as well as critical. In Section 2.6 there is a discussion of the difference between the mass functions of Lodieu et al. (2012) and Bouy et al. (2015), where a stronger conclusion needs to be drawn. In particular the effects of differing theoretical mass-luminosity relationships is given as a possible reason for the difference, but the candidate needs to find the functions used, and then calculate whether the differences are sufficient to explain the discrepancy in mass functions. Similarly, the effects on the mass functions of the differing distances used should be calculated. Finally, some overall comment about the differences in total masses derived for the cluster (Section 2.6.1) should be provided.

Factual issues with literature review.

On page 22, there is a discussion of the HIPPARCOS distance to the Pleiades. This needs to be framed in the context of the previous distances which had been derived using stellar models, which were at odds with that from the then new space-based parallaxes.

On page 26 the discussion of Pritchard (1884) misses the point. The quotation continues “and not simply that common motion of the whole which would necessarily arise simply from the translation of the solar system in space”. I.e. he did not detect proper motion, this should be made clear.

The final paragraph of page 64 needs a more technically sophisticated discussion, which should include equipartition of energy and violent relaxation as well as slower virialisation.

Defending the choice of methods.

As I stated in my commentary on the scientific quality of the thesis, I do not expect to agree with all the choices made in the analysis chain, and am also happy to accept that choices were made at the outset which the author thinks in retrospect could be improved. Indeed, I think such discoveries are an important part of the thesis. However, there are places where I think the reasons for the choices could be better laid out, and where a critical assessment is required in the light of the results.

A good example of this is the decision in Section 2.7.1 to use the $i-K_s$ colour index (CI) as the parameter for the models of the magnitudes. There are other choices which could have been made here, for example K_s , and so a discussion is needed of the advantages of this choice, or indeed the downside if the candidate believes his work shows there is now a better choice. The thesis should also discuss here the possibility of using mass as parameter, which I suspect does not work because it would require a model, but the possibility should be discussed.

I like the idea of modelling the sequence with splines (Section 3.3.4), and here a good justification is given, in terms of other methods tried. However, the candidate states that the sequence has to be continuous. This is not true in general for stellar isochrones, with their gradient discontinuities at key stages of their evolution reflecting structural changes. This is a well-known source of problems in interpolating isochrones. So, whilst I think it is desirable that the isochrones are continuous, I think a little more justification is required.

Section 3.3.4 describes how the model consists of a single-star sequence and an equal-mass unresolved binary sequence. What about unequal mass unresolved binaries? I do not want to see a complete re-analysis of the data here, but some discussion of the effects of missing out this sequence are required. I suspect it relates to the intrinsic spread which is derived for the sequence.

I don't completely follow the way that the spatial distribution on the sky has been fitted. I can understand how an extra term for the symmetry is introduced (Equation 3.51). What I'm unclear about (and I think this is just a problem of my understanding, but it needs explanation) is how Equation 3.52 is applied. Is the radius for each star measured and one instance of this equation applied for each star?

In the results section, the candidate states in the second paragraph of Section 4.8.1 that the mass-luminosity relationship used is for the 2MASS AB photometric system. I'm not sure what is going on here, the 2MASS system is nominally a Vega one not an AB system. The situation is complicated by the fact it takes quite a careful reading of Cohen, Wheaton & Megeath (2003, AJ 126, 1090) to understand the zero points, and the documentation for the Allard et al files is not explicit. A first step in understanding this may be to compare the mass-luminosity relationships derived from the Allard et al AB and Vega files. Maybe also comparing the data with the models might ensure a gross error has not been made.

Discussing the conclusions.

On page 112 (3rd paragraph) I do not follow why missing data preferentially affect the field stars. What matters here is the ratio of field density to cluster density, not the absolute value. This then questions the analysis in the following paragraph which suggests that this, along with the large proper motion uncertainties causes the discrepancy in membership probabilities. I suspect it is purely the PM uncertainties which cause the difference, but the candidate needs to give some numbers to show that this is (or is not) the case.

The statement at the end of Page 120 is worrying, since it implies that the priors have a significant effect on the results. I would like to see some ratios of priors to posteriors so we can see what their effect is.

The plots in Figure 4.19 are perturbing, in the sense that I do not understand why the Gaussian mixture model produces long wings to the posterior. These wings will have a tendency to include some high proper

motion objects. I would like the candidate to explore whether they are real or a product of the model. Do the simulations show anything similar? Why does the EMB sequence not show something similar?

In the discussion of the mass-luminosity relationship (Page 149) it is stated that “the mass-luminosity relation relies entirely on the current models of stellar atmospheres”, whereas in fact it is largely the structure models which matter. This statement should be corrected.

Figure 4.24 shows that the mass functions derived from the *J*, *H* and *K* bands are very different around a log mass of -1.3, but there is no discussion of this, and the thesis then moves on to using the K-band with no detail of why. A discussion of this is needed.

In the final paragraph of Chapter 4 no mention is made about systematic uncertainties, in particular in the mass-luminosity relationship, which might affect the comparison of clusters of different ages. The range of possible mass-luminosity relationships should be discussed, with a quantitative conclusion. Also, unless the candidate can suggest something, the comments about comparing two calculated datasets using Bayes’ theorem seems to be misplaced. Bayes’ theorem allows one to compare two models of data, not two datasets which should be consistent.

Minor changes requested.

The second equation in in section 1.1 should read

$$\xi(\log_{10}M) = \frac{dN}{d(\log_{10}M)}$$

(My main issue is the lack of the infinitesimal denominator, but also *N* needs to be italic, and various, as do various symbols in the equation above it.)

Third line from the bottom of page 64, astronomical -> astrometric.

On page 73 I think

“...plane of the sky *R*. Volume In this work we will assume that the analytical expression in Eq. 3.41 is correctly stated in terms of the number density and not of volume density” should read

“...plane of the sky *R*. In this work we will assume that the analytical expression in Eq. 3.41 is correctly stated in terms of the projected number density and not of projected mass density”

p(a) is nicely set up as a probability distribution in Section 3.1, but in Equation 3.15 it is used as a pure probability. The candidate should use the convention of capitalising *P* in these cases.

On page 104 the sentence “To do this, I compare the membership probabilities, summarised by the mode, recovered after inferring the model on two synthetic data sets of” seems to have gone mangled. I assume this means that a single synthetic data set was created with complete set of observables for all stars, and then a second was created from it, with some values masked as missing.